Color Scene Analysis

Mehmet Celenk
Ohio University, Department of Electrical and Computer Engineering
Stocker Center, Athens, Ohio 45701

Abstract

This paper describes a color scene analysis method for the object surfaces appearing in the noisy and imperfect images of natural scenes. It is developed based on the spatial and spectral grouping property of the human visual system. The uniformly colored surfaces are recognized by their monomodal 3-D color distributions and extracted in the spatial domain using the lightness and chromaticity network of the Munsell system. The textured image regions are identified by their irregular histogram distributions and isolated in the image plane using the Julesz connectivity detection rules. The method is applied to various color images corrupted by noise and degraded heavily by undersampling and low color-contrast imperfections. The method was able to detect all the uniformly colored and heavily textured object areas in these images.

1. Introduction

Scene analysis is the first essential step in the development of vision systems.[1] This process determines the simple structural characteristics of objects' surfaces using detailed image domain properties. It involves the separation of object surfaces from the background area in noisy and imperfect images of complex scenes involving irregular textures. Although the identification of object and surface boundaries comes naturally to a human observer, accurate scene analysis has proved difficult and complex. Achieving an adequate scene analysis result depends mainly on devising techniques to detect uniformity among the feature values of the picture points, and then isolating the areas of the picture exhibiting these uniformities. In this paper, we describe a method to find the visually distinct and contextually meaningful image areas corresponding to the uniformly colored or heavily textured object surfaces.

First, the uniformly colored object regions are recognized by their monomodal color distributions in the 3-D color space using lightness and chromaticity network of the Munsell system. The textured image regions are then identified by their irregular histogram distributions. A texture analysis scheme is developed based on Julesz's conjecture to obtain the structural scene characteristics of textured regions. Features extracted include the spatial and spectral information embodied in the textured areas. This enables the method to extract the textured object regions using the image dependent features.

The algorithm is applied to the four color images of the city skyline taken during winter. These images were corrupted by noise and degraded heavily by different imperfections (e.g., low resolution and very low color contrast). The method was able to detect all the uniformly colored or heavily textured object areas in these images.

Remaining of this paper is organized as follows. We first give the background information for selecting the color space for the method. Operation of the overall approach is then described by an algorithm. Analysis results of an under-sampled, low-contrast picture are then presented.

2. Background

A color image is usually given by three values at every pixel, which correspond to the R (red), G (green), and B (blue) tristimuli. For color scene analysis, it is desirable that the selected color features define a space possessing uniform metric.[2,3,4] The (L*,a*,b*) and (U*,V*,W*) color coordinate systems developed by the CIE (Commission Illumination d'Eclairage) in 1976[5] approximately satisfy this property. It has been shown in Ref.[6] that the former space gives better results than the latter in analyzing the color pictures. This system is obtained from the (R,G,B)-primary system by converting the (R,G,B) values into the (X,Y,Z)-nonphysical primary system[7]

$$X = 2.7690R + 1.7518G + 1.1300B \tag{1.1}$$
$$Y = 1.0000R + 4.5907G + 0.0601B \tag{1.2}$$
$$Z = 0.0000R + 0.0565G + 5.5943B \tag{1.3}$$

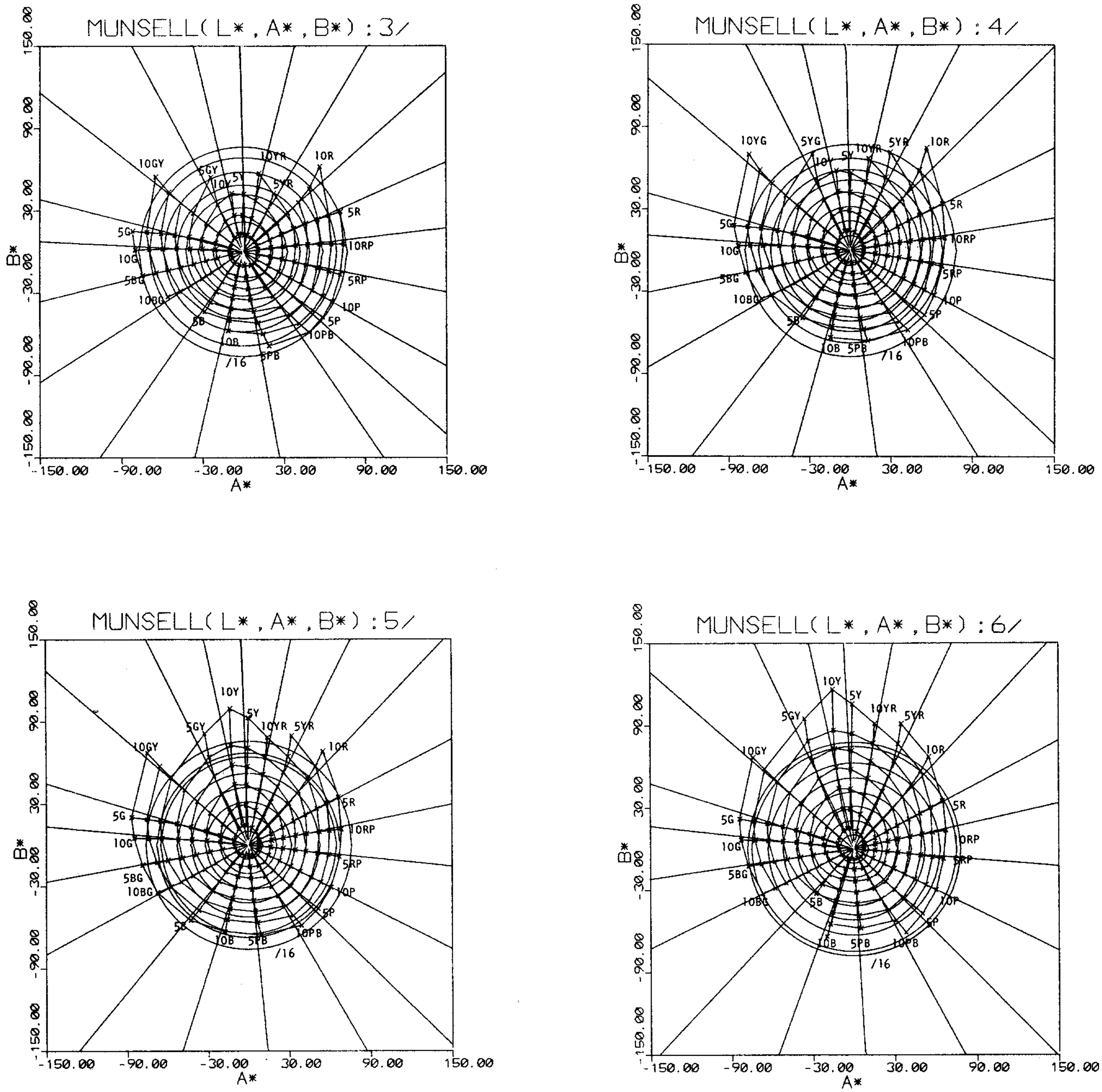
and applying a cube-root transformation to the (X,Y,Z) values:

$$L^* = 116[Y/Y_0]^{1/3} - 16 \qquad ,Y/Y_0 > 0.01 \tag{2.1}$$
$$a^* = 500[(X/X_0)^{1/3} - (Y/Y_0)^{1/3}] \qquad ,X/X_0 > 0.01 \tag{2.2}$$
$$b^* = 200[(Y/Y_0)^{1/3} - (Z/Z_0)^{1/3}] \qquad ,Z/Z_0 > 0.01 \tag{2.3}$$

where X_o, Y_o, and Z_o are the (X,Y,Z) values of the reference white. Here, they are selected 2^n-1 for n-bit image data representation. The cylindrical coordinates (L*, H°, C*)[8] of this space resemble the empirical Munsell color order system[9] and concur almost exactly with the accepted physiological model of color vision.[10] These coordinates, known as psychometric lightness, hue and chroma, are given by

$$L^* = L^* \tag{3.1}$$
$$H^\circ = \tan^{-1}(b^*/a^*) \tag{3.2}$$
$$C^* = (a^{*2} + b^{*2})^{1/2} \tag{3.3}$$

The loci of constant lightness, hue, and chroma describe the appearance of object colors under certain view and lighting conditions. Any horizontal section through the space would define a plane of constant lightness while any vertical plane originating at the achromatic L* axis would



MUNSELL(L* , A* , B*) : 3/
MUNSELL(L* , A* , B*) : 4/
MUNSELL(L* , A* , B*) : 5/
MUNSELL(L* , A* , B*) : 6/
B*
A*
150.00
90.00
30.00
-30.00
-90.00
-150.00
/16

Figure 1.

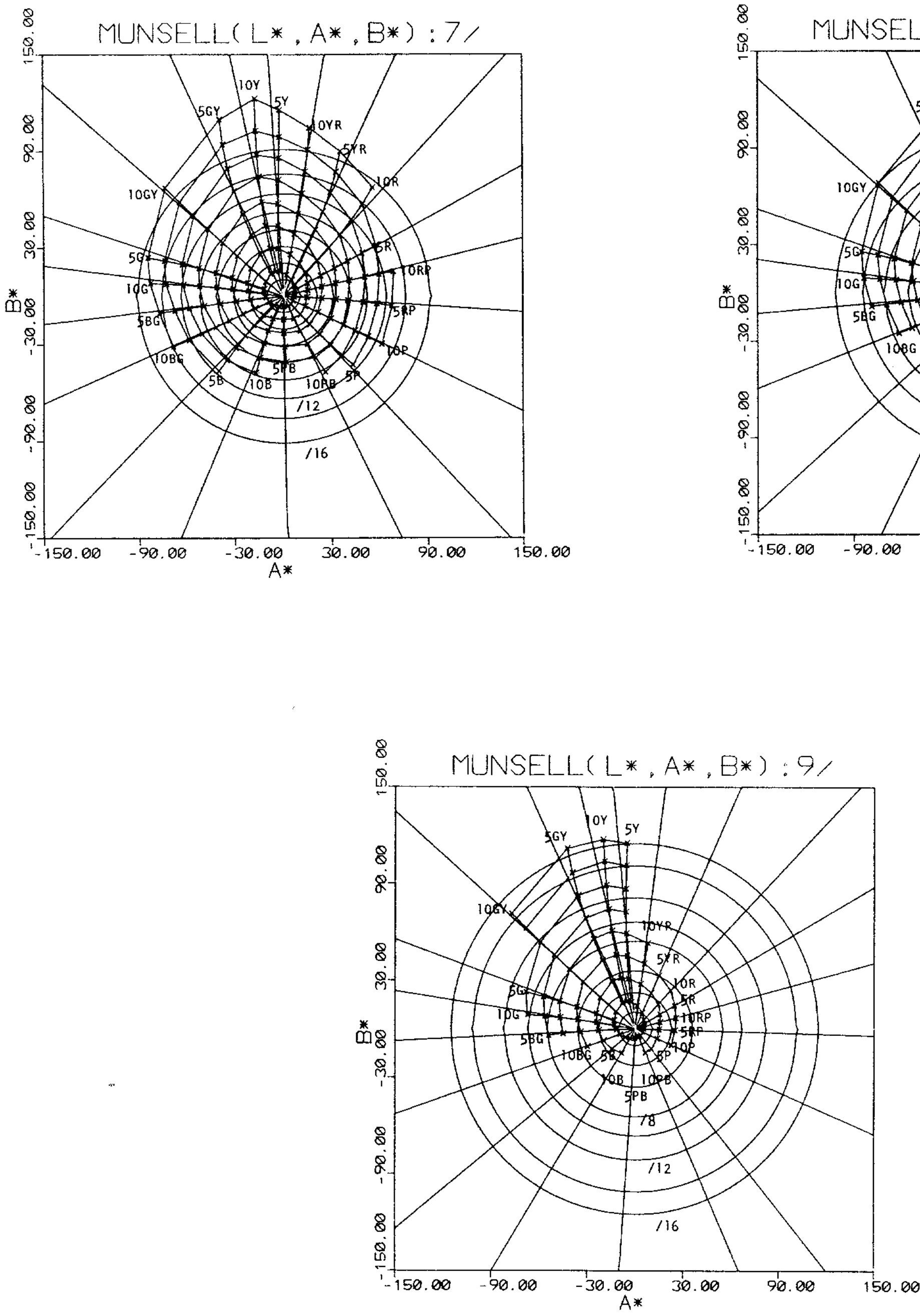


Figure 1, continued.

be a plane of constant hue. A cylindrical section concentric with the L* axis however would constitute a surface of constant saturation in which all the object colors have the same purity. In Fig.1, organization of the constant chromaticity network of the Munsell system are shown for twenty hues (varying from 5R to 10RP) at eight chroma values (increasing from /2 to /16) in the (a*,b*)-plane for the Munsell values of 3/ through 9/. The radial hue and concentric chroma loci are also approximated by radial lines and concentric circles in the least square sense as shown in the same figure.

3. Description of the Method

First, the uniformly colored object regions are recognized by their monomodal color distributions in the (L*,a*,b*) space. They are extracted using lightness and chromaticity network of the Munsell color order system. The textured image regions are then identified by their irregular histogram distributions. The textured surfaces are separated from the background area using the Julesz conjecture. Operation of the method may be described by the following algorithm:[12,13,14]

Step 1: Starting from the 1-D histograms of the line projections of the (L*,a*,b*) space, detect the most prominent color distribution and its neighbor(s) in some ranges of a particular color component.

Step 2: Based on this line information, project only a portion of the space onto the other two color coordinates and using their respective l-D histograms, determine the plane properties of the detected color distributions.

Step 3: Based on the extracted plane information, project only a part of the space onto the last color coordinate and using the respective 1-D histogram, find the space distributions of the detected modes.

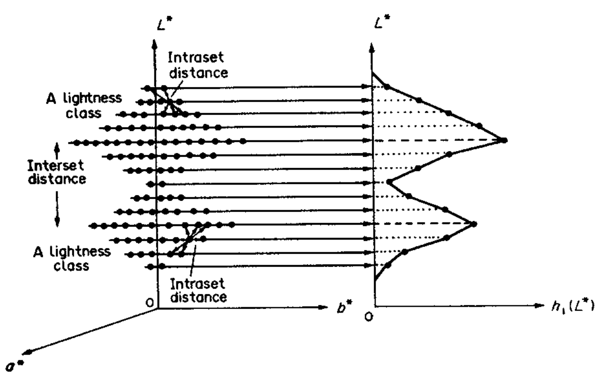
Step 4: Project the estimated color distributions of the best mode and its neighbor onto the line of Fisher discriminant for 1-D thresholding.

Step 5: Test the modality of 1-D histograms of remaining image points for a prominent color distribution. If a decisive 1-D peak exists, proceed with step 1; otherwise, extract an additional feature set and proceed with step 1.

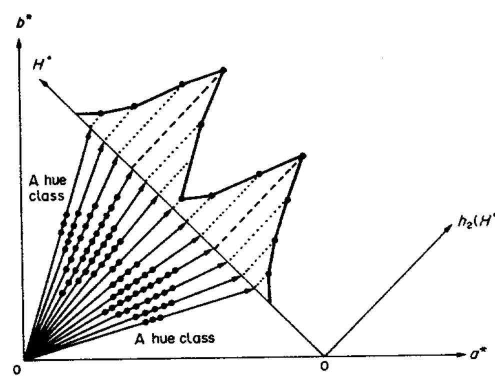
In the following sections, we first demonstrate the operation of the first three steps of the algorithm for analyzing the uniformly colored object surfaces. We then present the operation of step 5 for isolating the heavily textured regions or surfaces in the scene.

3.1. Uniformly Colored Surface Analysis

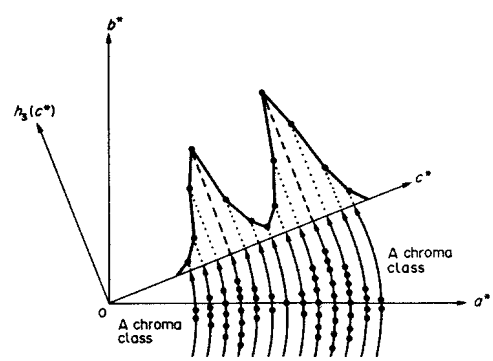
The Munsell network of colors is explored for estimating the 3-D color distributions of the uniformly colored object regions. For this purpose, the radial hue lines and the concentric chroma circles of Fig. 1 are used to estimate the object color distributions in some circular-cylindrical volume elements in the (L*,a*,b*) space. The underlying volume is bounded by two lightness planes, two hue planes, and two chroma cylinders. Figure 2 shows the process of forming such a decision element. It is clear from Fig. 2 that we need to determine the values of six loci to specify such a decision volume. This gives the nonparametric estimates of the object color distributions in the 3-D color space. The



(a) Bimodal histogram of different lightness classes.



(b) Bimodal histogram of different hue classes.



(c) Bimodal distributions of different lightness hue and chroma classes.

Figure 2.

goal here is to reduce the computational cost involved in forming 3-D decision surfaces. In addition, use of the 1-D histograms for detecting the 3-D color distributions makes the detection process computationally efficient. This is similar to the process described in Ref.[11], which relies on the use of elliptical and parabolic decision surfaces to enclose image clusters in the X-Y, X-I, and Y-I plane projections of the (X,Y,I) normalized color space. However, use of the chromaticity network results in simpler decision surfaces than those obtained in Ref.[11]. Furthermore, here the color distribution is sought in the 3-D space instead of its lower dimensional subspaces as proposed in Ref.[11]. This permits utilization of all the property values of object colors for the analysis, and inherently recognizes their respective cross correlation. This way, the region acceptance is not limited to the information available from one color component.

3.2. Textured Surface Analysis

After extracting all the uniformly colored regions, the picture is then analyzed for textured regions. For this purpose, the remaining part of the image is partitioned into the atomic regions of maximally 4-connected pixels of the similar color. The 4-connectivity between an image pixel (i,j) and its neighbor (m,n) in a 3×3 local window centered at (i,j) is defined by $|i - m| + |j - n| = 1$. For this partitioning, every unprocessed pixel is initially assumed to be an atomic region by assigning it a region number. Starting from the first row and first column, the color vector of each unclassified point is compared with that of its left and top neighbors. If a similarity between the pixel in process and any of its specified neighbors in a 3×3 local window is detected, then the region number of that element is modified to that of its similar neighbor. Here the similarity between an image pixel (i,j) and its left or top neighbor (m,n) is defined by

$$|L^*(i,j) - L^*(m,n)| < L_t \tag{4.1}$$
$$|H^\circ(i,j) - H^\circ(m,n)| < H_t^\circ \tag{4.2}$$
$$|H^\circ(i,j) - H^\circ(m,n)| > H_t^\circ \tag{4.3}$$
$$|C^*(i,j) - C^*(m,n)| < C_t \tag{4.4}$$

where L_t, H_t°, and C^t are the image dependent threshold values. The outlined region growing technique is repeated at every unprocessed point of the image.

After this initial partitioning, the algorithm computes the center of G_k of every atomic region a_k using

$$i_k = \left\{ \sum_i \sum_j |C(i,j)| \cdot i \right\} / \left\{ \sum_i \sum_j |C(i,j)| \right\} \tag{5.1}$$

$$j_k = \left\{ \sum_i \sum_j |C(i,j)| \cdot j \right\} / \left\{ \sum_i \sum_j |C(i,j)| \right\} \tag{5.2}$$

where $|C(i,j)|$ is the magnitude of the color vector of a point (i,j) in a_k. An $M \times N$ local window is then located at the center of the atomic region in process (see Fig.3). Within this window, eight fixed directions are defined from the center of that atomic region. On each of these directions, similarity between the atomic region in question and its neighbors is tested according to the similarity criteria given by Eqs. (4.1) through (4.4). If the continuity property is observed along these directions, the atomic region of interest is classified as being a part of a uniform region and its feature vector is computed in the corresponding direction(s). If the continuity search fails, then a possible periodic texture structure is sought to classify the atomic region as being a spectral primitive of a texture field and to compute its feature vector in the respective directions. The spatial

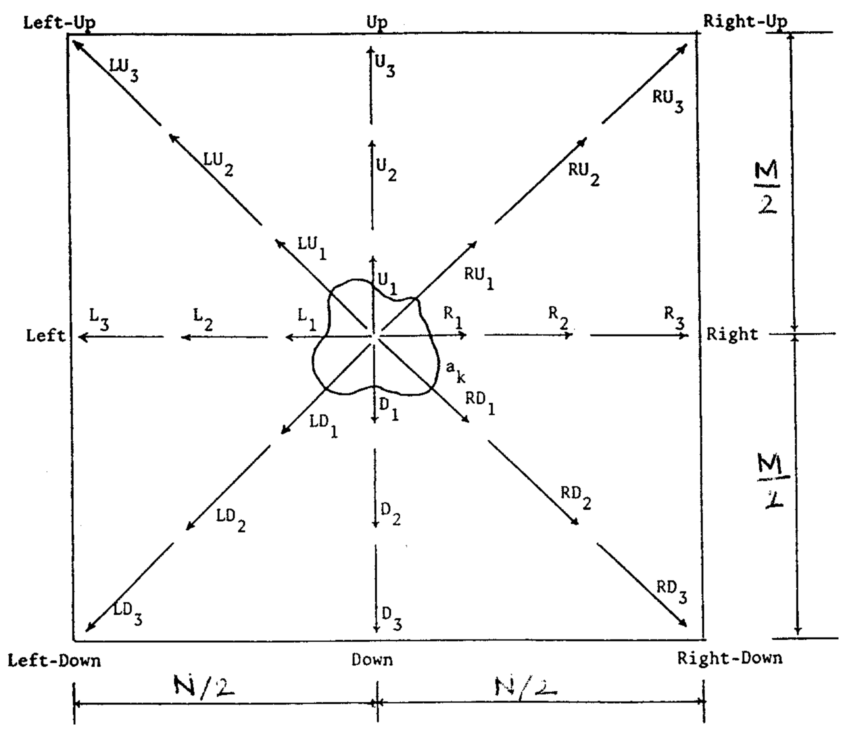


Figure 3.

primitives of a texture field are the atomic regions and the spectral primitives are the colors which satisfy the Julesz connectivity detection rules[15,16,17] for spontaneous discrimination of visual patterns. This conjecture states that the human eye has the capability of grouping the similar colors, but no grouping of dissimilar colors is possible. Figure 4 illustrates this property in the (x,y)-chromaticity plane. This property is observed in the spatial domain as follows: Suppose that we are given four different hue samples of the same luminance involving red (R), yellow (Y), green (G), and blue (B) colors. Assume that the color patches are created from the given hue samples and arranged in the spatial plane in two different view fields as shown below:

View field 1:		View field 2:	
Texture1	Texture 2	Texutre 1	Texture 2
RYRYRYRY	GBGBGBGB	RGRGRGRG	YBYBYBYB
YRYRYRYR	BGBGBGBG	GRGRGRGR	BYBYBYBY
RYRYRYRY	GBGBGBGB	RGRGRGRG	YBYBYBYB
YRYRYRYR	BGBGBGBG	GRGRGRGR	BYBYBYBY

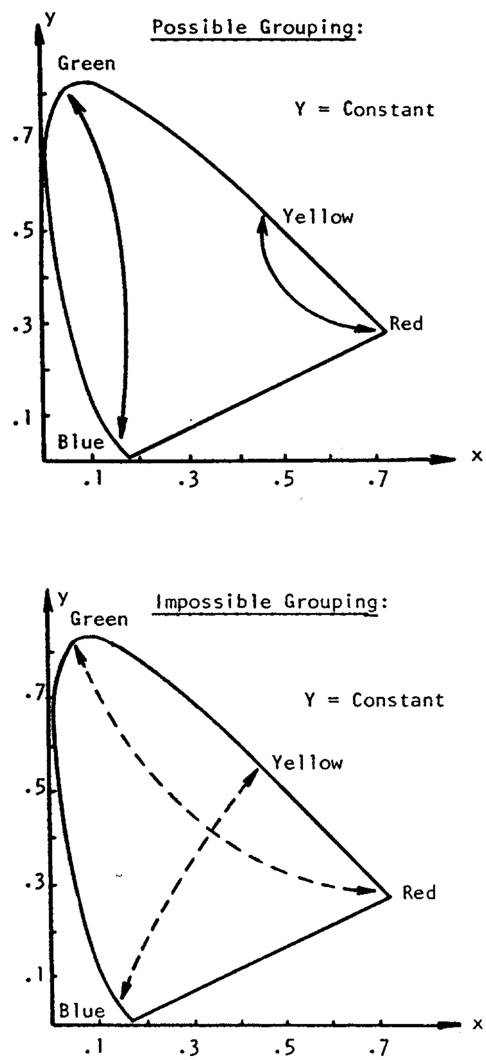


Figure 4.

In spontaneous vision, the human eye will group the given color patches as arranged in the view field 1. This suggests that the color transitions for a texture field should be sought by considering the atomic regions with least color deviation. Thus the spatial periodicity in this search is defined as the 3-step random color transition that takes place between adjacent atomic regions. It also implicitly satisfies the spectral property stated above. The following expression illustrates the 3-step color transition:

$$C_k \rightarrow C_l \rightarrow C_k \rightarrow C_l \qquad (6)$$

where C_k and C_l are the color patch vector of atomic regions of a texture field as shown in Fig. 5. An important property of this transition is that it is not color-contrast dependent and not a function of the structural properties (e.g., shape, size, orientation, etc.) of a textured region. If a color vector is modified by a constant or the spatial structure of a texture is changed, the above transition takes place as long as the difference between the consecutive color patches exceeds the threshold given above.

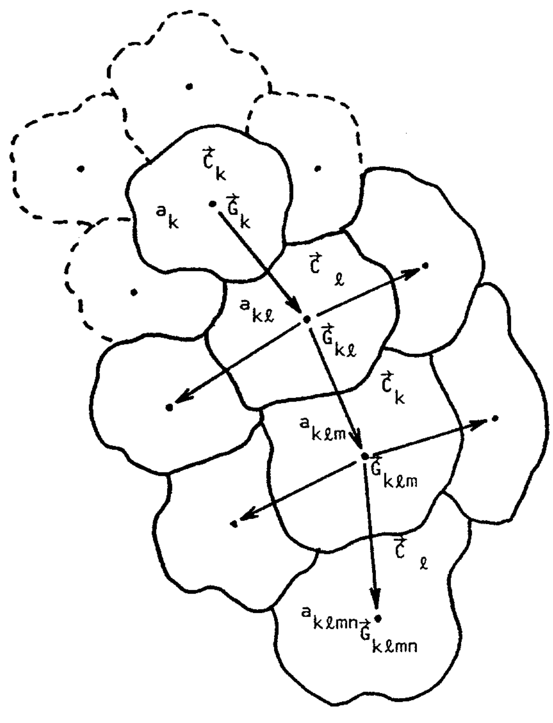


Figure 5.

In the absence of continuity and texture patterns, the atomic region is accepted as an isolated one and its feature assignment is made considering all of its touching neighbors.

Once an atomic region is classified as being a part of a visual pattern L (see Fig. 6) through the specified directions, then the feature vector f is assigned to that atomic region by averaging the color vectors of the atomic regions in that visual pattern as described below:

$$f = E\{C_k \cup C_l\} \qquad (7)$$

Here, C_k and C_l are the color vectors of the atomic region a_k and its neighbors in the determined directions in L, E is the expected value operator, and U represents the union of color vectors. Since the E operator simulates the grouping property of the human eye in L, the region L is then analyzed as a uniformly colored area with respect to f.

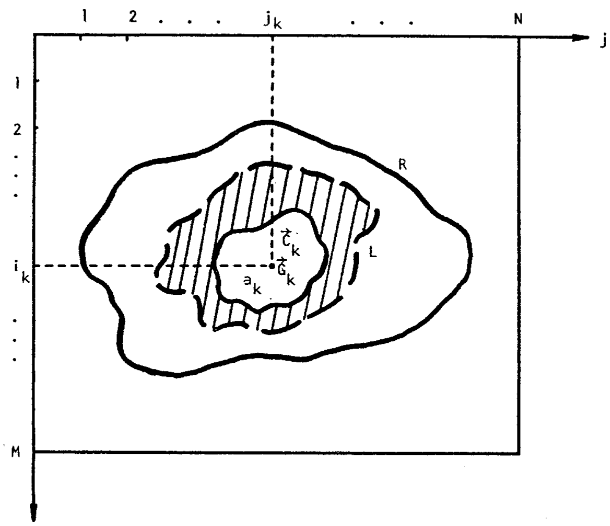


Figure 6.

4. Experimental Results

The color scene analysis technique presented in this paper was applied to several low contrast color images of the skyline of Pittsburgh city taken in winter. The spectral contents of the images are limited mainly to the bright colors. This property is observed mostly on the lightness and chroma distributions, which are concentrated in the lower values of chroma and the higher values of lightness (usually C*<30 and L*>60), respectively. The original images are given with the 8 bits/pixel R,G,B specification in 580 × 700 grid. The pictures were resampled and their sizes are reduced to 193 × 232 by a factor of 3 in each dimension. This under sampling created various image imperfections such as missing elements and noise.

Four images were analyzed in the computer implementation of the method. Only the analysis results of one of these pictures are presented here. The original image (in reduced form) is given in Fig. 7 and the processing results are shown in Fig. 8.

5. Conclusions and Further Research Topics

In this paper, a color scene analysis technique has been presented and its use in color picture segmentation has been described. The procedure described here does not use any a priori information about the scene domain or impose any constraints on the color distributions. It also extracts the features most useful for the picture being processed. With these properties, the technique is not affected by substantial variations of input scenes.

Further research is needed on the following issues. The link between the presented method and the sensor characteristics and their physical models requires further research. Its operation needs to be improved for analyzing scenes involving rough textures, shadows and highlights, inconsistent color appearance, etc. Operation of the method should be tested on oddly shaped color distribution (e.g., linear, long, elongated, etc.).

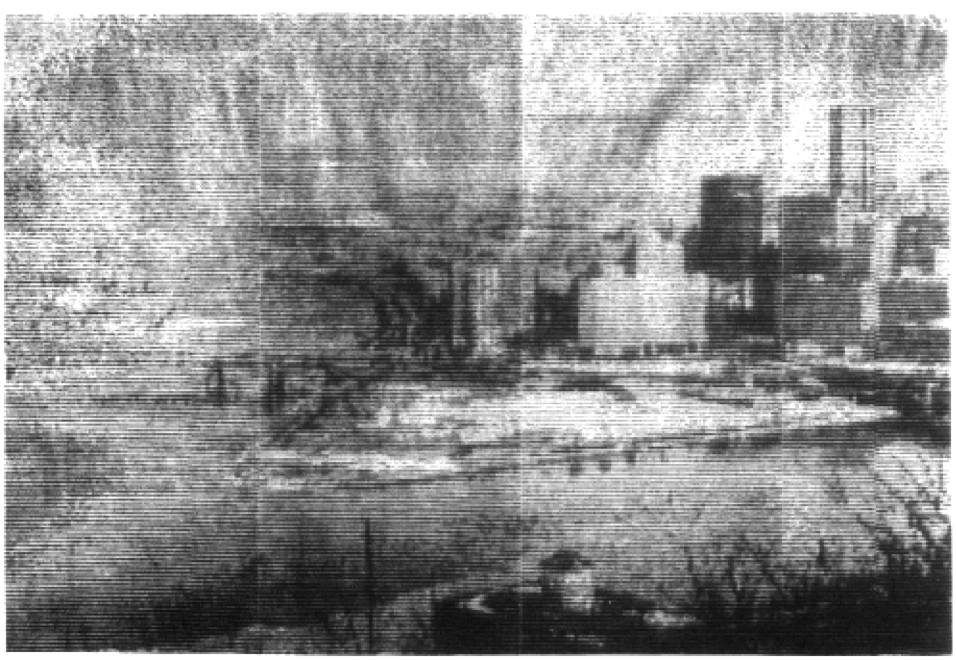
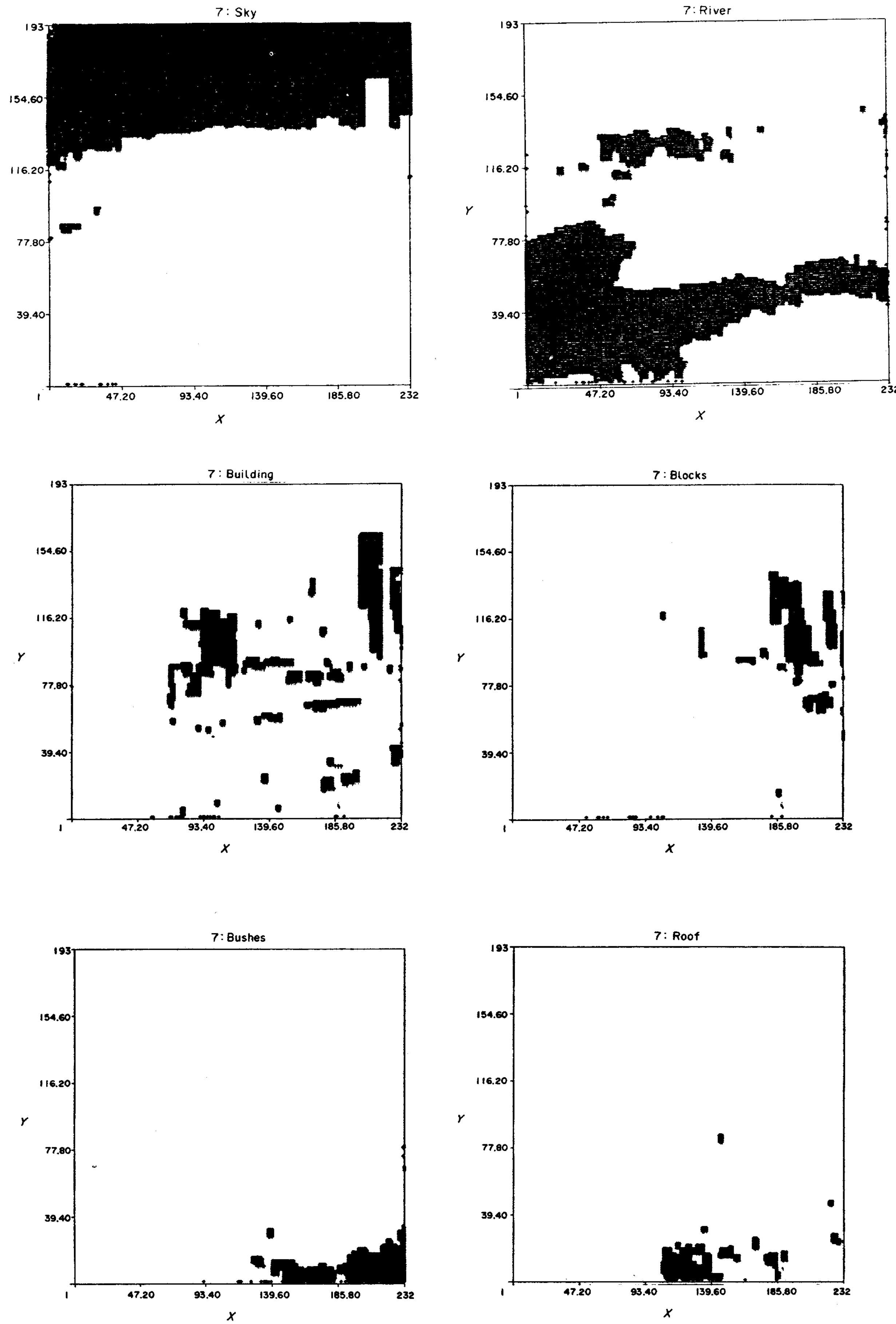


Figure 7.


7: Sky
7: River
7: Building
7: Blocks
7: Bushes
7: Roof
193
154.60
116.20
77.80
39.40
Y
1
47.20
93.40
139.60
185.80
232
X
Figure 8.

6. References

1. R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, Vol. **I**, Addison Wesley, 1993.
2. F. Bumbaca and K. C. Smith, "Design and implementation of colour vision model for computer vision applications," *CVGIP* **39**, pp. 226-245, 1987.
3. G. B. Coleman, Image Segmentation by Clustering, *USCIPI Report* **750**, Univ. of Southern California, 1977.
4. S. Tominaga, "Color image segmentation using three perceptual attributes," *Proc. Conf. Computer Vision and Pattern Recognition*, pp.628-630, 1986.
5. A.L. Robertson, "The CIE 1976 color difference formulae," *Color Research and Application* Vol. **2**, No.1, pp. 7-11, Spring 1977.
6. Y. Ohta, T. Kanade and T. Sakai, "Color information for region segmentation," *CGIP* **13**, pp. 222-241, 1980.
7. W. T. Wintringham, "Color television and colorimetry," *Proc. IRE*, Vol. **39**, pp. 1135-1172, 1951.
8. K. McLaren, "The development of the CIE (L*,a*,b*)-uniform color space," *JSDC*, pp. 338-341, 1976.
9. G. Wyszecki and W. S. Stiles, *Color Science*, Wiley, 1967.
10. R. H. Wallis, Film Recording of Digital Color Images, *USCIPI Report* **570**, Univ. of Southern California, 1975.
11. M. Ali, W. N. Martin and J. K. Aggarwal, Color-based computer analysis of aerial photographs, *CGIP* **9**, pp. 282-293, 1979.
12. M. Celenk, "A color clustering technique for image segmentation," *CVGIP* **52**, pp. l45-170, 1990.
13. M. Celenk, "An adaptive machine learning algorithm for color image analysis and processing," *Int. J. Robotics & Computer Integrated Manufacturing,* Vol. **4**, No. 3/4, pp. 403-412, 1988.
14. M. Celenk, "Colour image segmentation by clustering," *IEEE Proceedings*, Vol.**138**, No. 5, pp. 368-376, Sep. 1991.
15. B. Julesz, "Visual pattern discrimination," *IRE Trans. Information Theory*, pp. 84-92, 1962.
16. B. Julesz, "Texture and visual perception," *Scientific American*, pp. 38-48, Feb. 1965.
17. B. Julesz, "Experiments in the visual perception of texture," *Scientific American*, pp. 34-43, April 1975.

published previously in SPIE, Vol. 2179, page 407